JOURNAL OF MULTIVARIATE ANALYSIS **20**, 129–142 (1986)

Limits of Translation Invariant Experiments

ARNOLD JANSSEN

Universität-Gesamthochschule Siegen, Siegen, West Germany

Communicated by J. Pfanzagl

In this article we study the behaviour of translation invariant experiments. The main result establishes necessary and sufficient conditions for the local approximation of a sequence of suitably normalized product experiments by Gaussian shifts with respect to the weak convergence. Moreover, it turns out that the local asymptotically normal (LAN) condition with uniform remainders with respect to the local parameter can be characterized in terms of the local behaviour of the Hellinger distances. This result depends on the fact that a convergent sequence E_n is always equicontinuous if E_n is convergent. The result answers a question which has been posed in the author's article about "the convergence of almost regular statistical experiments to Gaussian shifts" (In *Proceedings, 4th Pannonian Sympos. Math. Statist.*, Bad Tatzmannsdorf, 1983). The proofs rely on a theorem showing that binary experiments can be treated by positive definite functions.

The local approximation of a normalized sequence of experiments $E_n = (\Omega^n, \mathfrak{A}^n, (P_{\delta_n \theta})_{\theta \in \mathbb{R}^k})$ by Gaussian shifts is one of the most important areas in asymptotic statistics. In the past various conditions for the validity of a LAN expansion (local asymptotically normal) of the corresponding likelihood processes were published, cf. LeCam [9], Hájek [3], Ibragimov and Has'minskii [4]. The results are based on the fundamental paper of LeCam [9]. Stronger approximations were considered by LeCam [8], Michel and Pfanzagl [13], Milbrodt [14, 15].

In the literature there exists a number of known examples—called almost regular experiments—where the LAN condition is fulfilled but the standard assumptions, the L^2-differentiability of LeCam, are violated. In a preceding paper [6] the author made an attempt to attack the problem for translation invariant experiments $E = (\Omega, \mathfrak{A}, (P_\theta)_{\theta \in \mathbb{R}^k})$. In this connection a partial answer was given. If, roughly speaking, the square of the Hellinger

Received February 4, 1985; revised March 6, 1985.

AMS 1980 subject classifications: 62E20, 62A05.

Key words and phrases: local asymptotic normality, limit experiments, translation invariance, stability, positive definite functions.

0047-259X/86 $3.00

distance $t \to d^2(P_0, P_{tx})$ is regularly varying at 0 (for all $x \in \mathbb{R}^k - \{0\}$) with the exponent $p = 2$, then there exists a Gaussian shift which is a weak accumulation point of E_n.

This paper aims to show that the condition mentioned above is necessary and sufficient for the LAN condition. Moreover, it turns out that for continuous translation invariant experiments the convergence of E_n is equivalent to the equicontinuity of the sequence. Thus the LAN condition implies a LAN condition with uniform remainders w.r.t. the local parameter. The arguments can also be used to prove equicontinuity of general sequences $U_{\delta_n} E^n$ which yield stable limit experiments in the sense of Strasser [18].

The proof are based on a transform for binary experiments which introduces the connection to positive definite functions. The transform has various applications. For example, $t \to 1 - d^2(P_0, P_t)$ is seen to be positive definite which applies to further problems.

1. Preliminaries

In this section the notation is introduced and some definitions of LeCam [12] are recalled. Let $E = (\Omega, \mathfrak{A}, (P_\theta)_{\theta \in \Theta})$ be a statistical experiment. Then E^n denotes the nth product experiment of E consisting of the product measures P_θ^n. If $\Theta = \{1, 2\}$ then we write E in the condensed form $\{P_1, P_2\}$. Denote by

$$H_E(z) = \int \prod_{i \in I} \left(\frac{dP_i}{d\mu_I}\right)^{z_i} d\mu_I$$

the Hellinger transform of E which is defined on the set $\{z = (z_i)_{i \in I} \in [0, 1]^I : \sum_{i \in I} z_i = 1,\ I \text{ finite}\}$, $P_i \ll \mu_I$ for $i \in I \subset \Theta$. Note that $H_{E^n} = (H_E)^n$. If E, F are experiments for the same parameter space Θ then $H_E = H_F$ iff $E \sim F$ in LeCam's sense [12].

Let dP_t/dP_s denote the Radon–Nikodym density of the absolutely continuous part of P_t with respect to P_s, cf. [19]. Then the weak convergence of classes of experiments (with respect to $\sim$) is defined by the weak convergence of all finite dimensional marginal distributions of the log-likelihood ratio process

$$\mathscr{L}((\log(dP_t/dP_s))_{t \in I} \mid P_s) \qquad \text{for all finite subsets } I \subset \Theta, \quad s \in I.$$

By definition

$$d(P, Q) = \left(\frac{1}{2} \int \left(\left(\frac{dP}{d(P+Q)}\right)^{1/2} - \left(\frac{dQ}{d(P+Q)}\right)^{1/2}\right)^2 d(P+Q)\right)^{1/2}$$

is the Hellinger metric and $\|\cdot\|$ is the norm of total variation of measures. It is known that $H_{\{P,Q\}}(\frac{1}{2}, \frac{1}{2}) = 1 - d^2(P, Q)$ and $2d^2(P, Q) \leqslant \|P - Q\| \leqslant 2\sqrt{2}\, d(P, Q)$.

If T is a linear space then an experiment $E = (\Omega, \mathfrak{A}, (P_t)_{t \in T})$ is said to be translation invariant if $E \sim (\Omega, \mathfrak{A}, (P_{t+h})_{t \in T})$ for all $h \in T$. It is continuous if $t \to P_t$ is $\|\cdot\|$-continuous, provided T is a metrizable real vector space. Define

$$U_a E = (\Omega, \mathfrak{A}, (P_{at})_{t \in T}) \qquad \text{for} \quad a \geqslant 0, a \in \mathbb{R}.$$

An experiment E on $\mathbb{R}^k$ is called a Gaussian shift if E is equivalent to the experiment $(\mathbb{R}^k, \mathscr{L}^k, (N(\Gamma t, \Gamma))_{t \in \mathbb{R}^k})$, where $N(a, \Gamma)$ denotes the normal distribution with mean a and covariance matrix Γ on $\mathbb{R}^k$. A Gaussian shift has the Hellinger transform $H_{\{P_0, P_t\}}(z, 1-z) = \exp((z^2 - z)\, \frac{1}{2} t' \Gamma t)$. A positive function L on the interval $(0, \delta)$ is slowly varying if $L(ts)\, L(t)^{-1} \to 1$ for $t \to 0$ and all $s > 0$.

2. Main Results

(2.1) Lemma. *Let $E_n = (\Omega_n, \mathfrak{A}_n, (P_{n,\theta})_{\theta \in \mathbb{R}^k})$ be a sequence of translation invariant continuous experiments such that $E_n \to E$, $E = (\Omega, \mathfrak{A}, (P_\theta)_{\theta \in \mathbb{R}^k})$, is weakly convergent. Then the subsequent assertions are equivalent.*

(i) *E_n is equicontinuous, i.e., $\theta \to P_{n,\theta}$ is equicontinuous with respect to $\|\cdot\|$.*

(ii) *E is continuous.*

The implication (i) $\Rightarrow$ (ii) is well known and valid without restrictions while the converse assertion heavily relies on the translation invariance. The non-trivial part applies to the following situation.

(2.2) Corollary. *Let E be a translation invariant continuous experiment for an arbitrary normed vector space T. If $U_{\delta_n} E^n$ tends weakly to a continuous experiment F for a sequence $\delta_n \downarrow 0$ of real numbers then F is stable,*[1] *i.e., $U_{n^{1/p}} F \sim F^n$ for some $p > 0$ and each $n \in \mathbb{N}$,* cf. [18].

In case of translation invariant experiments the equicontinuity of $U_{\delta_n} E^n$ always holds on finite dimensional subspaces of T. Thus (2.2) sharpens Theorem (2.9) of Strasser [18] if the experiments are translation invariant.

Next we consider a Gaussian shift F with covariance Γ on $T = \mathbb{R}^k$. Then F is continuous, translation invariant, and stable with the index of stability $p = 2$. A sequence $U_{\delta_n} E^n$ is said to be locally asymptotically normal (LAN) if $U_{\delta_n} E^n$ tends weakly to a Gaussian shift F. The concept of LAN sequences

[1] F is called stable or scale invariant.

goes back to LeCam [9], see also Hájek [3], Ibragimov and Has'minskii [4], and Strasser [19, (80.6)]. It is our intension to give a version of the LAN condition which is uniform on compact sets w.r.t. the remainders of the local parameter. On the other hand, necessary and sufficient conditions in terms of the Hellinger distance are proved.

(2.3) THEOREM. *Let* $E=(\Omega, \mathfrak{A}, (P_t)_{t\in\mathbb{R}^k})$ *be a continuous translation invariant experiment and* $\Gamma=(\gamma_{ij})$ a $k\times k$ *covariance matrix such that* $\gamma_{11}>0$. *Then the subsequent assertions equivalent*:

(i) *There exists a sequence of positive real numbers* $\delta_n\to 0$ *such that* $U_{\delta_n}E^n$ *tends weakly to the Gaussian shift F with covariance* Γ.

(ii) *There exists a function L on* $(0,\infty)$ *being slowly varying at* 0 *and positive in a neighbourhood of* 0 *such that*

$$\frac{d^2(P_0, P_{tx})}{t^2L(t)}\to\frac{x'\Gamma x}{\gamma_{11}} \qquad \text{if} \quad t\to 0+$$

for all $x\in\mathbb{R}^k$, $t\in(0,\infty)$.

(iii) *There exists a sequence of positive real numbers* $\delta_n\to 0$ *such that the likelihood ratio admits the following expansion*:

$$dP^n_{\delta_n t}/dP^n_0=\exp(t'X_n-\tfrac{1}{2}t'\Gamma t+R_n(t)).$$

(1) $\mathscr{L}(X_n\mid P^n_0)\to N(0,\Gamma)$ *weakly.*

(2) $\lim_{n\to\infty}\sup_{t\in K}P^n_0\{|R_n(t)|\geq\varepsilon\}=0$

for every $\varepsilon>0$ *and all compact subsets* $K\subset\mathbb{R}^k$.

The results are proved in Section 5.

(2.4) *Remarks.* (i) The theorem improves a result of the author [6, Theorem 4]. In this paper it is pointed out that (2.3)(ii) is necessary for LAN sequences of translation invariant experiments. Conversely it has been proved that the Gaussian shift F is a weak accumulation point of $U_{\delta_n}E^n$ if (ii) is fulfilled. Note that (2.3) together with the results of [6] implies that $(\Omega^n, \mathfrak{A}^n, (\otimes_{j=1}^n P_{d_{nj}t})_{t\in\mathbb{R}^k})$ tends to a Gaussian shift if $\sum_{j=1}^n d_{nj}^2L(|d_{nj}|)$ converges to $a\neq 0$ and $\max_{j\leq n}|d_{nj}|\to 0$.

(ii) The LAN-condition only depends on the local structure of the Hellinger transforms at 0 in case of translation invariant experiments. It is only necessary to compute the local behaviour of the Hellinger transforms which has been done in various cases for location experiments. In view of applications compare with the discussion of [6].

(iii) The condition (2.3)(iii) is referred to be a LAN-condition uniform on compacts w.r.t. the remainders in connection with the local parameter which is equivalent to the equicontinuity of $U_{\delta_n} E^n$. In standard regular situations (when $L(t) \to a > 0$ for $t \to 0$) the uniformity is well known, see LeCam [9] and Ibragimov and Has'minskii [4]. The LAN-condition in a stronger setting is used to establish global asymptotic normality, see Milbrodt [14, 15]. In these papers the problem is to find a global approximation of the experiments by suitable exponential families which are asymptotically normal, see Michel and Pfanzagl [13] and earlier papers of LeCam [8].

(iv) It is known that in the situation of (2.2) the equicontinuity can be deduced if the Hellinger distance satisfies a two-sided Lipschitz condition locally at 0 like $c_1 |t|^a \leqslant d(P_0, P_t) \leqslant c_2 |t|^a$, $0 < c_1 < c_2$. In the translation invariant case only the convergence and the continuity of the experiments are needed.

3. Binary Experiments

The proofs depend on the subsequent results for binary experiments. They have been studied by LeCam [10, Chaps. 2, 3] (see also Strasser [19, Chap. IV]). Suppose that $F = (\Omega, \mathfrak{A}, (P, Q))$ is a binary experiment. Let μ be a σ-finite measure on Ω dominating F. Then we define the transform

$$\mathscr{T}_F(y) := \int_{\{dP/d\mu > 0, dQ/d\mu > 0\}} \left(\frac{dP}{d\mu}\right)^{1/2 + iy} \left(\frac{dQ}{d\mu}\right)^{1/2 - iy} d\mu$$

for $y \in \mathbb{R}$, where i denotes the complex unit.

Recall that a complex function $\varphi: T \to \mathbb{C}$ is positive definite if for all n-tuples $t_1, ..., t_n$ of elements of T, the $n \times n$ matrix $(\varphi(t_i - t_j))$ is positive semidefinite with respect to $\mathbb{C} \times \mathbb{C}$.

(3.1) *Remark.* (1) $\mathscr{T}_F$ is independent of the dominating measure μ.

(2) $|\mathscr{T}_F(y)| \leqslant \mathscr{T}_F(0) = H_F(\frac{1}{2}, \frac{1}{2}) = 1 - d^2(P, Q)$.

(3) $\mathscr{T}_F(y) = \int (dP/dQ)^{1/2 + iy}\, dQ = \int (dQ/dP)^{1/2 - iy}\, dP$.

(4) $y \to \mathscr{T}_F(y)$ is a continuous positive definite function.

(5) $\mathscr{T}_F(y) = \int_{\mathbb{R}} \exp((\frac{1}{2} - iy)\, x)\, d\mu_1(x)$, where $\mu_1 = \mathscr{L}(\log(dQ/dP) \mid P)$ denotes the log-likelihood distribution of P on $[-\infty, \infty)$.

The roofs are left to the reader. Note that

$$\frac{dP}{dQ} = \frac{dP}{d\mu} \left(\frac{dQ}{d\mu}\right)^{-1} \qquad \text{on the set} \qquad \left\{\frac{dP}{d\mu} > 0, \frac{dQ}{d\mu} > 0\right\}.$$

The transform $\mathscr{T}_F$ has the following meaning. The log-likelihood distribution μ_1 has finite Laplace transform $\omega(z)=\int_{\mathbb{R}} \exp(zx)\,d\mu_1(x)$ on the strip $\{z\in\mathbb{C}: 0<\operatorname{Re} z<1\}$ of the complex plane. The Hellinger transform $H_F(t, 1-t)$ equals $\omega(1-t)$ for real $t\in(0,1)$, whereas the transform $\mathscr{T}_F(y)$ equals the Laplace transform $\omega(\frac{1}{2}-iy)$ on the line $\frac{1}{2}-iy$, $y\in\mathbb{R}$. There is another interpretation of $\mathscr{T}_F(y)$. According to (5), $\mathscr{T}_F(-y)$ is the Fourier transform of the measure having the μ_1-density $x\to\exp(\frac{1}{2}x)$. The total mass of this measure is $H_F(\frac{1}{2},\frac{1}{2})$. Moreover $\mathscr{T}_F=\mathscr{T}_E$ if $F\sim E$.

(3.2) LEMMA. *Let F_n be binary experiments, $n\in\mathbb{N}\cup\{0\}$. Then*

(a) $\mathscr{T}_{F_1}(y)\cdot\mathscr{T}_{F_2}(y)=\mathscr{T}_{F_1\otimes F_2}(y)$, *where $F_1\otimes F_2$ denotes the product experiment of F_1 and F_2.*

(b) $\mathscr{T}_{F_1}=\mathscr{T}_{F_2}$ *iff* $F_1\sim F_2$.

(c) *Assume that $\mathscr{T}_{F_n}(y)\to h(y)$ is convergent for all $y\in\mathbb{R}$. Then there exists a binary experiment F such that $\mathscr{T}_F=h$ and F_n tends weakly to F.*

(d) *The following assertions are equivalent:*

(i) $F_n\to F_0$ *weakly.*

(ii) $\mathscr{T}_{F_n}(y)\to\mathscr{T}_{F_0}(y)$ *for all* $y\in\mathbb{R}$.

(iii) *Assertion* (ii) *holds uniformly on compact subsets of* $\mathbb{R}$.

Proof. Define $F_n=(\Omega_n, \mathfrak{A}_n, (P_n, Q_n))$.

(a) It is well known that the log-likelihood distribution of a product experiment

$$\mathscr{L}\left(\log\frac{dQ_1\otimes Q_2}{dP_1\otimes P_2}\,\middle|\, P_1\otimes P_2\right)$$

coincides with the convolution product of

$$\mathscr{L}\left(\log\frac{dQ_1}{dP_1}\,\middle|\, P_1\right)*\mathscr{L}\left(\log\frac{dQ_2}{dP_2}\,\middle|\, P_2\right),$$

which is defined on $[-\infty,\infty)$, cf. [7, (9.3), (9.21)]. Clearly, the Laplace transform of a convolution product is the product of the Laplace transforms.

(b) Assume that $\mathscr{T}_{F_1}=\mathscr{T}_{F_2}$. In view of (3.1) the uniqueness theorem for Fourier transforms implies that $\mathscr{L}(\log(dQ_1/dP_1)\mid P_1)=\mathscr{L}(\log(dQ_2/dP_2)\mid P_2)$. This fact can be proved by considering first the measures with the density $\exp(\frac{1}{2}x)$. Hence $H_{F_1}=H_{F_2}$, giving $F_1\sim F_2$.

(c), (d) First we show that (i) implies (ii). Recall that the weak convergence of F_n is equivalent to the weak convergence of the standard measures

$$\mathscr{L}\left(\left(\frac{dP_n}{d(P_n+Q_n)}, \frac{dQ_n}{d(P_n+Q_n)}\right) \,\middle|\, P_n+Q_n\right)$$
$$\to \left(\left(\frac{dP}{d(P+Q)}, \frac{dQ}{d(P+Q)}\right) \,\middle|\, P+Q\right)$$

on the simplex $S_2 = \{(x_1, x_2) \in [0, 1]^2 \colon x_1 + x_2 = 1\}$. Since

$$\mathscr{T}_F(y) = \int x_1^{1/2+iy} x_2^{1/2-iy} \, d\mathscr{L}\left(\left(\frac{dP}{d(P+Q)}, \frac{dQ}{d(P+Q)}\right) \,\middle|\, P+Q\right)(x_1, x_2),$$

the result follows.

Since $(F_n)_n$ is weakly sequentially compact, each weak accumulation point F of $(F_n)_n$ has the transform $\mathscr{T}_F = h$ and (c) follows. It is well known that a sequence of continuous positive definite functions $\mathscr{T}_{F_n}$ converges uniformly on compact sets if pointwise convergence is assumed. ▮

In view of applications the main theorem for the transform $\mathscr{T}$ is proved.

(3.3) THEOREM. *Let $E = (\Omega, \mathfrak{A}, (P_t)_{t \in T})$ be a translation invariant experiment for a vector space T. Then for fixed $y \in \mathbb{R}$ the function $\varphi_y \colon T \to \mathbb{C}$, $\varphi_y(t) := \mathscr{T}_{\{P_0, P_t\}}(y)$, is positive definite on T.*

Proof. Suppose that $z_1, \ldots, z_n \in \mathbb{C}$, $t_1, \ldots, t_n \in T$. Choose a σ-finite measure μ dominating $P_{t_1}, \ldots, P_{t_n}, f_{t_i} := dP_{t_i}/d\mu$. Taking $\{P_{t_j}, P_{t_i}\} \sim \{P_0, P_{t_i - t_j}\}$ into account we arrive at

$$\begin{aligned}\sum_{i,j=1}^{n} z_i \bar{z}_j \varphi_y(t_i - t_j) &= \sum_{i,j=1}^{n} z_i \bar{z}_j \mathscr{T}_{\{P_{t_j}, P_{t_i}\}}(y) \\ &= \int \sum_{i,j=1}^{n} z_i \bar{z}_j f_{t_j}^{1/2+iy} f_{t_i}^{1/2-iy} \, d\mu \\ &= \int \left| \sum_{i,j=1}^{n} z_i f_{t_i}^{1/2-iy} \right|^2 d\mu \geqslant 0\end{aligned}$$

and the result is proved, $\bar{z} := \operatorname{Re} z - \operatorname{Im} z$. ▮

Inserting $y = 0$ we remark that $t \to 1 - d^2(P_0, P_t)$ is positive definite.

(3.4) LEMMA. *If E is translation invariant for $T = \mathbb{R}^k$ then the following assertions are equivalent:*

(i) *$t \to P_t$ is $\|\cdot\|$-continuous.*

(ii) φ_y *is a continuous function for fixed* $y \in \mathbb{R}$.

(iii) φ_0 *is continuous in* 0.

Proof. (i) $\Rightarrow$ (ii) Assumption (i) implies the convergence of the binary experiments $\{P_0, P_t\}$ to $\{P_0, P_s\}$ for $t \to s$. According to (3.2) $\varphi_y(t) \to \varphi_y(s)$ for $t \to s$ proves the continuity of φ_y at s.

(iii) $\Rightarrow$ (i) Note $\varphi_0(t) = 1 - d^2(P_0, P_t)$. Therefore $t \to P_t$ is continuous in 0. The translation invariance proves the statement. ▮

(3.5) EXAMPLE.. Let $E = (\mathbb{R}^k, \mathscr{L}^k, (N(\Gamma t, \Gamma))_{t \in \mathbb{R}^k}$ be a Gaussian shift with covariance Γ. Then

$$\varphi_y(t) = \exp(-\tfrac{1}{8} t'\Gamma t - \tfrac{1}{2} y^2 t'\Gamma t).$$

The following Lemma is well known. A proof is contained in the article of Milbrodt [16, (3.6), (4)].

(3.6) LEMMA. *Let E_n be an equicontinuous family of experiments for a separable metric parameter space Θ. Then $(E_n)_n$ is weakly sequentially compact.*

4. MISCELLANEOUS RESULTS

In this section let E denote a continuous translation invariant experiment for a normed parameter space T (3.3). Define the positive definite function φ by $\varphi(t) = 1 - d^2(P_0, P_t)$ according to (3.3). E is said to be infinitely divisible if for every $n \in \mathbb{N}$ there exists a nth root of E, cf. [7]. By the same definition infinitely divisible probability measures are introduced.

(4.1) *Remark.* If E is infinitely divisible for $T = \mathbb{R}^k$ then there is a uniquely determined symmetric infinitely divisible probability measure μ on $\mathbb{R}^k$ having the Fourier transform $\hat{\mu} = \varphi$.

Theorem (3.3) has further immediate consequences which illustrate the connection between stable distributions and stable experiments.

(4.2) *Remark.* Let E be a translation invariant and stable experiment with index $p > 0$, i.e., $U_{t^{1/p}} E = E^t$, c.f. [18].

(a) If $T = \mathbb{R}^k$ then φ is the Fourier transform of a symmetric stable distribution on $\mathbb{R}^k$ with the index p of stability. φ is known to be equal to $\varphi(tx) = \exp(-|t|^p \cdot f(x))$, $x \in S := \{y \in \mathbb{R}^k : |y| = 1\}$, $t \in \mathbb{R}$, where $f: S \to [0, \infty)$ is a continuous function such that $f(s) = f(-s)$.

(b) The index of stability p is not greater than 2.

Note that $\varphi(y)^t = \varphi(yt^{1/p})$ in case of stable experiments. Assertion (b) is well known for stable distributions. Thus (a) implies (b). For experiments the result (b) was proved in [5] by a different proof.

Let

$$r(y) = \limsup_{t \to 0} \frac{\log d(P_0, P_{ty})}{\log t}$$

denote the degree of regularity with respect to the direction $y \in T$. Then Pflug [17] shows that r is bounded by 1 for location experiments. In general this is a result for positive definite functions.

(4.3) *Remark.* $r(y) \leqslant 1$.

Proof. If $r(y) > 1$ is assumed then there is a sequence $t_n \to 0$ such that

$$\frac{1-\varphi(t_n)}{t_n^2} = \frac{d^2(P_0, P_{t_n y})}{t_n^2} \to 0.$$

By [1, (5.3)], $\varphi(ty) \equiv 1$ follows for all t, which is impossible. (Note that [1, (5.3)] applies to a sequence $t_n \to 0$). ∎

(4.4) Lemma. *For every infinitely divisible continuous translation invariant experiment E there is a continuous translation invariant Gaussian experiment $F = (X, \mathscr{L}, (Q_t)_{t \in T})$ having the same φ function on T and therefore the measures of F have the same Hellinger distances $d(P_0, P_t) = d(Q_0, Q_t)$.*

Proof. It is sufficient to prove (4.4) for $T = \mathbb{R}^k$. By Schoenberg's theorem [1, (7.8)]

$$a(t) := -\log \varphi(t) = -\log(1 - d^2(P_0, P_t))$$

is a negative definite function. This means that

$$K(s, t) := 4(a(s) + a(t) - a(s-t))$$

is positive semi-definite. Note that $1 - d^2(P_s, P_t) = \varphi(s-t)$. It is well known that the Gaussian experiment F with covariance K has the same Hellinger distance as E, cf. [7; or 18, (3.3)]. ∎

(4.5) *Remark.* Let E be a translation invariant and stable Gaussian experiment for $T = \mathbb{R}^k$. Then there exists a continuous symmetric function $f: S \to [0, \infty)$ on the sphere S such that the covariance K equals

$$K(x, y) = 4\left(|x|^p f\left(\frac{x}{|x|}\right) + |y|^p f\left(\frac{y}{|y|}\right) - |x-y|^p f\left(\frac{x-y}{|x-y|}\right)\right)$$

for $x \neq y$, $K(x, x) = 8\,|x|^p f(x/|x|)$, and $f(x) = -\log \varphi(x)$.

In case $k=1$ the result is due to Strasser [18, (4.1)]. Suppose that E is Gaussian with the index of stability $p=2$. Then φ belongs to a normal distribution and $f(x)=\frac{1}{2}x'\Gamma x$ follows. Therefore E is a Gaussian shift and Theorem (4.3) of [18] is proved.

5. Proofs of the Main Results

Proof of Lemma (2.1). (ii) $\Rightarrow$ (i) Choose $\psi_n(t)=\mathscr{T}_{\{P_{n,0},P_{n,t}\}}(0)$ according to (3.3). Then ψ_n is a sequence of continuous positive definite functions on $\mathbb{R}^k$ being normed by $\psi_n(0)=1$. By (3.2) $\psi_n(t)\to\mathscr{T}_{\{P_0,P_t\}}(0)=:\psi(t)$, which is a consequence of the convergence of binary subexperiments. Note that ψ is also positive definite and continuous. The continuity theorem for Fourier transforms shows $|\psi_n(t)-\psi(t)|\to 0$ uniformly on compact subsets $K\subset\mathbb{R}^k$. Choose a compact neighbourhood K of 0 in $\mathbb{R}^k$ such that $|1-\psi(t)|<\varepsilon/2$ on K. Then there is a positive integer n_0 such that

$$\begin{aligned} d^2(P_{n,0},P_{n,t}) = 1-\psi_n(t) &\leqslant |\psi_n(t)-\psi(t)| \\ &\quad + |1-\psi(t)| < \varepsilon \end{aligned}$$

for all $t\in K$ and all $n\geqslant n_0$. Thus E_n is equicontinuous in 0 which is sufficient to prove Lemma (2.1), taking the translation invariance into account. The conclusion (i) $\Rightarrow$ (ii) is well known, cf. Strasser [18]. ∎

Proof of Corollary (2.2). By (2.1) the restriction of $U_{\delta_n}E^n_{|T'}$ on finite dimensional subspaces $T'\subset T$ is equicontinuous. Thus $F_{|T'}$ is stable, cf. Strasser [18, Theorem (2.9)], which proves the result for F. ∎

The proof of Theorem (2.3) relies on the following result.

(5.1) Lemma. *Let $E=(\Omega,\mathfrak{A},(P_t)_{t\in\mathbb{R}^k})$ be an infinitely divisible translation invariant experiment such that $H_{\{P_0,P_t\}}(\frac{1}{2},\frac{1}{2})=\exp(-\frac{1}{8}t'\Gamma t)$, where $\Gamma\neq 0$ is a covariance matrix. Then E is a Gaussian shift with covariance Γ.*

Proof. The assumptions imply that E is continuous.

1. Suppose that E is symmetric which means that $\{P_0,P_t\}\sim\{P_t,P_0\}$ is valid for all t. Note that

$$\limsup_{h\to 0}\frac{d^2(P_0,P_h)}{|h|^2}<\infty.$$

Therefore LeCam's results imply that $U_{n^{-1/2}}E^n$ tends weakly to the Gaussian shift G with the desired covariance, cf.[11; 8, (3)].

On the other hand, choose φ_y according to (3.3). An extension of (4.1) shows that there is a symmetric infinitely divisible probability measure μ_y on $\mathbb{R}^k$ such that $\hat{\mu}_y = \varphi_y$. By (3.2) (d) and (3.5)

$$\varphi_y(t/\sqrt{n})^n \to \exp(-\tfrac{1}{8}t'\Gamma t - \tfrac{1}{2}y^2 t'\Gamma t) \tag{I}$$

follows. Put $h_t\colon \mathbb{R}^k \to \mathbb{R}$, $x \to t'x$, $\nu_y := \mathscr{L}(h_t \mid \mu_y)$ and fix $t \neq 0$. If $t'\Gamma t = 0$ then $\nu_y = \varepsilon_0$ since $\{P_0, P_t\}$ is trivial. Otherwise by (I) ν_y belongs to the domain of normal attraction and ν_y has finite second moment, namely, $\frac{1}{4}t'\Gamma t + y^2 t'\Gamma t$, cf. Gnedenko and Kolmogorov [2, p. 181].

By Kolmogorov's formula for the Fourier transforms of infinitely divisible distributions (cf. [2, p. 85]) there exist a constant $\sigma_y^2 = \sigma_{y,t}^2 \geqslant 0$ and a finite measure $M_y = M_{y,t}$ on $\mathbb{R}\backslash\{0\}$ such that the Fourier transform equals

$$\hat{\nu}_y(s) = \varphi_y(st) = \exp\left(-\frac{\sigma_y^2}{2}s^2 + \int_{\mathbb{R}-\{0\}} (\cos(xs) - 1)\frac{1}{x^2}\, dM_y(x)\right).$$

Let $\nu_y^{1/n}$ denote the nth root of ν_y and define $T_a(x) = ax$. Then we prove

$$\mathscr{L}(T_{\sqrt{n}} \mid \nu_y^{1/n}) \to N(0, \sigma_y^2). \tag{II}$$

Note that the left-hand side of (II) has the Fourier transform

$$\exp\left(-\frac{\sigma_y^2}{2}s^2 + \int_{\mathbb{R}-\{0\}} \frac{1}{n}\left(\cos\,(x\sqrt{n}\,s) - 1\,\frac{1}{x^2}\right) dM_y(x)\right).$$

The integrand tends pointwise to 0 and it is uniformly bounded, (cf. below). Hence (II) follows from Lebesgue's theorem. The inequality $1 - \cos x \leqslant x^2$ implies

$$\frac{1}{n}(1 - \cos(x\sqrt{n}\,s))\frac{1}{x^2} \leqslant s^2.$$

Next, consider the sequence $U_{\sqrt{n}}E^{1/n}$, where $E^{1/n}$ is the nth root of E. Each of these experiments has the same Hellinger distance as E. Thus our sequence is equicontinuous. Applying (3.6), there exists a subsequence $n(j)$ such that $U_{\sqrt{n(j)}}E^{1/n(j)}$ tends weakly to some F which is continuous and translation invariant. Let ψ_y be the function defined in (3.3) for F. Then

$$\varphi_y(\sqrt{n(j)}\,st)^{1/n(j)} \to \psi_y(st) \qquad \text{for all } s \in \mathbb{R}.$$

Hence (II) implies $\psi_y(st) = \exp(-(\sigma_y^2/2)\,s^2)$.

Clearly, F also satisfies the assumptions of (5.1). Applying LeCam's result once more the convergence of $U_{n^{-1/2}}F^n \to G$ implies

$$\psi_y(st) = \psi_y(st/\sqrt{n})^n \to \exp(-\tfrac{1}{8}s^2t'\Gamma t - \tfrac{1}{2}y^2s^2t'\Gamma t)$$

and equality follows.

If we compare the different expressions for ψ_y, $\sigma_y^2 = \frac{1}{4}t'\Gamma t + y^2t'\Gamma$ follows.

By Kolmogorov's formula there exists a symmetric probability measure ρ_y such that $\nu_y = N(0, \sigma_y^2) * \rho_y$. Since the second moment of ν_y and σ_y^2 coincide ρ_y must be degenerated, $\rho_y = \varepsilon_0$. Hence $\nu_y = N(0, \sigma_y^2)$ and $\varphi_y(t) = \exp(-\frac{1}{8}t'\Gamma t - \frac{1}{2}y^2t'\Gamma t)$.

Thus all binary subexperiments of E are Gaussian in the sense of [7]; compare with (3.2)(b) and (3.5). Since each infinitely divisible experiment is Gaussian if the binary experiments have this property the covariance kernel of E equals $(t, s) \to t'\Gamma s$ and the result is proved, cf. [7].

2. The general case. Put $\tilde{E} = (\Omega, \mathfrak{A}, (P_{-t})_{t \in \mathbb{R}^k})$. Then $E_1 = U_{2^{-1/2}}E \otimes U_{2^{-1/2}}\tilde{E}$ is symmetric and case 1 applies to E_1. Therefore $U_{2^{-1/2}}E$ is Gaussian since E is infinitely divisible. Now [7, (2.9)] yields the covariance structure of the Gaussian shift with the matrix Γ. ∎

Proof of Theorem (2.3). In order to give a clear proof some results of [6] have to be recalled. First note that for every $0 < c < 1$,

$$H_{\{P_0^n, P_{\delta_n}^n\}}(\tfrac{1}{2}, \tfrac{1}{2}) = (1 - d^2(P_0, P_{\delta_n}))^n \to c$$

implies $nd^2(P_0, P_{\delta_n}) \to -\log c$ and vice versa.

(ii) ⇒ (i) Choose $\delta_n = \inf\{s \in [0, \infty) : d^2(P_0, P_{se_1}) = \gamma_{11}/8n\}$ for sufficiently large n, cf. [6, Lemma 9]. e_1 denotes the first unit vector in $\mathbb{R}^k$. Hence

$$nd^2(P_0, P_{\delta_n x}) = \frac{\gamma_{11}}{8} \frac{d^2(P_0, P_{\delta_n x})}{\delta_n^2 L(\delta_n)} \delta_n^2 \frac{L(\delta_n)}{d^2(P_0, P_{\delta_n e_1})}$$

tends to $\frac{1}{8}x'\Gamma x$ for each x by assumption (ii) and

$$\lim_{n \to \infty} H_{\{P_0^n P_{\delta_n x}^n\}}(\tfrac{1}{2}, \tfrac{1}{2}) \to \exp(-\tfrac{1}{8}x'\Gamma x)$$

follows. Let now G be any weak accumulation point of $U_{\delta_n}E^n$. Then G is infinitely divisible since $U_{\delta_n}E^n$ arises from an infinitesimal triangular array, cf. [7]. Now Lemma (5.1) can be applied to G which shows that G coincides with the Gaussian shift F. Since the set of equivalence classes of experiments is weakly compact the convergence of $U_{\delta_n}E^n$ is proved.

(i) ⇒ (ii) Define L on $(0, \infty)$ by $d^2(P_0, P_{te_1}) = t^2L(t)$. Then

Lemma 10 of [6] proves that L is positive localy at 0 and slowly varying. The result was proved in [6] for $x'\Gamma x \neq 0$. In the general case the following arguments are used:

$$(1 - d^2)P_0,\ P_{\delta_n tx}))^n \to \exp(-\tfrac{1}{8}t^2 x'\Gamma x)$$

uniformly in t on compact sets.

Assume $t_k \to 0$, $t_k > 0$. Choose $n(k) \in \mathbb{N}$ such that

$$\delta_{n(k)} \leqslant t_k \leqslant \delta_{n(k)-1} \qquad \text{or} \qquad \delta_{n(k)-1} \leqslant t_k \leqslant \delta_{n(k)}.$$

Then

$$n(k)\, d^2(P_0,\ P_{t_k x}) \to \tfrac{1}{8} x'\Gamma x$$

follows since $\delta_{n(k)}/\delta_{n(k)-1} \to 1$.

The same arguments can be applied to e_1 and (ii) is proved. The equivalence of (i) and (iii) is well known if we take the equicontinuity of $U_{\delta_n} E^n$ into account (2.1) and compare, for example, with Strasser [19, (80.6), (80.13)]. ∎

REFERENCES

[1] BERG, CH., AND FORST, G. (1975). Potential theory on locally compact Abelian groups. *Ergeb. Math. Grenzgeb.* Vol. 87, Springer-Verlag, Berlin/New York.

[2] GNEDENCKO, B. V., AND KOLMOGOROV, A. N. (1968). *Limit Distributions for Sums of independent Random Variables.* Addison–Wesley, Reading, Mass.

[3] HÁJEK, J. (1972). Local asymptotic minimax and admissibility in estimation. In *Proc. 6th Berkely Sympos. on Math. Statist. and Prob.* Vol.1, pp. 175–194. Univ. of California Press, Berkely, Calif.

[4] IBRAGIMOV, I. A., AND R. Z. HAS'MINSKII, (1981). *Statistical Estimation.* Springer-Verlag, New York/Berlin.

[5] JANSSEN, A. (1986). Scale invariant exponential families and one-sided test problems. *Statist. Decisions,* in press.

[6] JANSSENS, A. (1985). The convergence of almost regular statistical experiments to Gaussian shifts. In *Proceedings, 4th Pannonian Sympos. Math. Statist.,* Bad Tatzmannsdorf, 1983 (F. Konecny *et al.,* Eds.), pp. 181–204.

[7] JANSSEN, A., MILBRODT, H., AND STRASSER, H. (1985). *Infinitely Divisible Statistical Experiments. Lecture Notes Statist.* Vol. 27, Springer-Verlag, New York/Berlin.

[8] LECAM, L. (1956). On the asymptotic theory of estimating and testing hypotheses. In *Proc. 3rd Berkeley Sympos. on Math. Statist. and Prob.*-I, pp. 129–156.Univ. Califirnia Press Berkeley, Calif.

[9] LECAM, L. (1960). Locally asymptotically normal families of distributions. *Univ. California Publ. Statist.* **3** 37–98.

[10] LE CAM, L. (1969). *Théorie asymptotique de la décision statistique.* Les presses de l'Université de Montréal, Montréal.

[11] LECAM, L. (1970). On the assumptions used to prove asymptotic normality of maximum likelihood estimates. *Ann Math. Statist.* **41** 802–828.

[12] LeCam, L. (1974). *Notes on Asymptotic Methods in Statistical Decision Theory.* Centre de recherches Mathématiques, Université de Montréal, Montréal.
[13] Michel, R., and Pfanzagl, J. (1970). Asymptotic normality. Metrika **16** 188–205.
[14] Milbrodt, H. (1981). *Asymptotic Normality of Families of Probability Measures.* Dissertation, University of Bayreuth, unpublished.
[15] Milbrodt, H. (1983). Global asymptotic normality. *Statist. Decisions* **1** 401–425.
[16] Milbrodt, H. (1985). Limits of experiments associated with sampling plans. *Z. Wahrsch. verw. Gebiete* **70** 131–156.
[17] Pflug, G. Ch. (1984). The degree of regularity is not estimable. *Statist. Decisions. Suppl.* **1** 61–67.
[18] Strasser, H. (1982). Stability and translation invariance of statistical experiments. *Probab. Math. Statist.* **5** 1–20.
[19] Strasser, H. (1985). *Mathematical Theory of Statistics.* De Gruyter Studies in Mathem. 7, Berlin-New York.